Chapter 2
Covariability in Abundance among Index Stocks of Columbia River Spring/Summer Chinook Salmon

Louis W. Botsford, University of California, Davis, CA
Charles M. Paulsen, Paulsen Environmental Research, Ltd.

July 7, 1997

Abstract

Although a number of spawning populations of spring/summer run chinook salmon in the Columbia River basin have declined, a search for covariability among different locations using several different combinations of spawner and recruitment data indicates no basin-wide covariability. There is, however, significant covariability among index populations within the three main sub-basins: the Snake River, the Mid Columbia and the John Day River. This covariability was much stronger and more consistent in data reflecting survival (i.e., the natural logarithm of recruits per spawner, the residuals from a fit to the Ricker stock-recruitment model) than in data reflecting abundance (i.e., spawning escapement). We also tested a measure of survival that did not require knowing the age structure of spawners, the ratio of spawners in one year to spawners four years earlier, and found that they gave similar results. Because intra-series correlation was substantial and varied with data type, we accounted for intra-series correlation in judging the significance of correlations. To reduce the errors involved in computing the effective degrees of freedom, we computed a generic effective degrees of freedom for each of several data types.

Introduction

Covariability between abundance at different locations is a valuable source of information regarding the dynamic structure and physical forcing of spatially distributed populations (e.g., Campbell and Mohn, 1983, Cohen et al. 1991, Elner and Campbell 1991, Koslow et al. 1987). Statistical description of the observed covariability over space provides the basis for formulating hypotheses that can then be tested through further analyses and specific measurements or experiments. Such analyses have employed a variety of methods and types of data. Here we describe a simple approach to description of covariability among stocks distributed over space, in which we take special care to account for the effects of covariability within series on the significance of covariability among series. The specific application is to spring/summer runs of chinook salmon (*Oncorhynchus tshawytscha*) in the Columbia River basin.

A number of stocks spring/summer chinook salmon on the Columbia River have declined in the past 20 years, and several of them are listed as threatened under the US Endangered Species Act. Management to reverse these declines requires quantitative description of the cause(s) of the decline. Description of causes and evaluation of the effectiveness of various management approaches are the primary goals of the Program for the Analysis and Testing of Hypotheses (PATH). PATH has taken a multi-pronged approach to these questions using a variety of methods and data sources. The analysis described here is an initial exploratory part of that process. Here we test the hypothesis that there has been a similar trend in the state indicators for spring/summer run chinook salmon populations throughout the Columbia River basin. The management implications of the hypothesis are: if all stocks (listed and otherwise) show similar patterns in escapement, recruitment, or other state variables, then some common forcing function (e.g., ocean conditions) would be the most likely candidate for the (joint) decline. On the other hand, if there are systematic differences among races, deems, sub-basins or stocks, then causal mechanisms on smaller spatial scales are more likely to be associated with the observed historical patterns.

Detecting Covariability

Covariability between time series can be detected by a variety of statistical methods (e.g., regression, various forms of nonparametric correlation). The various methods of estimating covariability are subject to a number of common problems: intra-series correlation, multiplicity of tests, use of proxy variables, introduction of covariability in pre-processing and others (Kope and Botsford 1990, Botsford and Brittnacher 1992). Here we use one of the most common descriptors of covariability, the Pearson product-moment correlation coefficient. Using this statistic to detect covariability between time series involves computing the correlation between the series, then testing to determine whether that value is statistically significant, i.e., whether it is large enough that there is a low probability that it arose from chance alone. We account for problems that arise from intraseries correlation to determine the relative level of confidence we can have in each specific results, but we do not deal with the other problems which are more or less common to all of the computed correlations. We view the results as identification of pattern, rather than a proof of causality. We use a probability level of 0.05, and assume the correlation coefficient has a Gaussian distribution when it is near zero.

A number of different expressions have been used to estimate the variance of the estimated correlation coefficient in the presence of intraseries correlation (Bayley and Hammersley 1946, Chelton 1983, 1984, Drinkwater and Myers 1987, Kope and Botsford 1990, Botsford and Brittnacher 1992, Botsford and Wainwright, in prep.). Here we use the general expression from Botsford and Wainwright (in prep.)

$$\mathrm{var}[R_{xy}(k)] = \Big\{ \frac{1}{N-|k|} \sum_{v=1-N+|k|}^{N-|k|-1} \left[1-\frac{|v|}{N-|k|}\right] [r_{xx}(v)\, r_{yy}(v) + r_{xy}(v+k)\, r_{xy}(k-v)]$$

$$- \frac{2r_{xy}(k)}{N} \sum_{v=1-N+|k|}^{N-1} \left[1-\frac{m}{N-|k|}\right] [r_{xy}(v)\, r_{yy}(v-k) + r_{xx}(v)\, r_{xy}(k-v)]$$

$$+ \frac{r^2_{xy}(k)}{2N} \sum_{v=1-N}^{N-1} \left[1 - \frac{|v|}{N}\right] [r2y_y(v \;+\; 2r2x_y(v \;\; \} \qquad (1)$$

$$\text{where } m = \begin{cases} -j & j < 0 \\ j-|k| & j > |k| \\ 0 & 0 \_ j \_ |k| \end{cases}$$

where N is the length of the time series and the $\rho_{xy}(t)$ is the true correlation between the two time series, x and y at lag t. Many of the ways of estimating this quantity from the series themselves involve substituting estimated values of correlations for the true correlations in various simplifications of this expression (Bayley and Hammersley 1946, Chelton 1983, 1984, Drinkwater and Myers 1987, Kope and Botsford 1990, Botsford and Brittnacher 1992, Botsford and Wainwright, in prep.).

When the null hypothesis is no correlation between the series (i.e., $\rho_{xy} = 0$), only the first term [i.e., the one involving $r_{xx}(v)r_{yy}(v)$] is nonzero (Kope and Botsford 1990). The expression involving only that term can be used to estimate the variance of the estimated correlation coefficient, or the expression can be rearranged to express the effective number of degrees of freedom in the time series, i.e., the effective sample size (cf., Bayley and Hammersley 1946),

$$\frac{1}{N^*} = \frac{1}{N} + \frac{2}{N} \sum_{j=1}^{N} \frac{N - j}{N} \rho_{xx_{(j)}} \rho_{yy_{(j)}} \qquad (2)$$

Note that when there is no auto-correlation in either of the series, the effective number of degrees of freedom is the length of the series, N.

There are several important consequences of this dependence of the standard error of the estimate of a correlation coefficient on intra-series correlation. From the above expressions it is apparent that as the degree of intraseries correlation increases, increasingly positive correlation between points at various lags will lead to a larger variance and a lower number of degrees of freedom. If this dependence is ignored, i.e., if the intra-series correlation is not accounted for, the frequency at which a significant correlation would be falsely detected (i.e., a Type I error) increases as the intra-series correlation increases. For example, as time series vary from a rapidly changing series of independent points to a slowly changing series such as a moving average of 4 adjacent independent points, the frequency of false detection of a significant result varies from the specified value of 0.05 to a frequency of 0.26. To avoid this increase in the number of spurious correlations identified as real correlations, one would account for intra-series correlation using equation (1) or (2).

The effect of this correction can be seen by examining the dependence of the value of correlation coefficient required to maintain a fixed error level on intra-series correlation as reflected by the number of independent degrees of freedom. A plot of the value of estimated correlation coefficient required to be significant at the 0.05 level versus the effective number of degrees of freedom of the series (Figure 2) shows that as the effective number of degrees of freedom decreases, the value of correlation required for a significant result increases substantially.

Because the correction that must be made for *intraseries* correlation leads to a higher required level of *interseries* correlation, it limits the number of true relationships we can detect to only the strongest. The implications of the above results for the detection of relationships between time series population for different populations follow from the relationship in Fig. 2. To correct for the intra-series correlation in slowly varying time series, we require a higher value of correlation for statistical significance. Thus, a certain level of actual correlation between two series (e.g., 0.5) would be detectable (i.e., could be identified as significant) in rapidly varying time series (e.g., with 16 degrees of freedom or more, for 0.5), but not in slowly varying series, with fewer degrees of freedom.

For very slowly varying time series, it is extremely unlikely that real covariability between series will be detected. For example, the number of degrees of freedom in linear trends is approximately 2 (i.e., the number of parameters required to express a straight line), hence covariability between two time trends would require an estimated correlation coefficient near 1.0 which is very unlikely given the amount of measurement error and other confounding noise typically associated with these series. For this reason, variability on such slow time scales is virtually undetectable. Linear trends reflect the slowest time scale, but similar, lesser, effects would occur as the data varied more rapidly. For example, series dominated by half a period of a cyclic fluctuation would have only a few more degrees of freedom than a linear trend, and so on. Covariability in time series involving linear trends or up to one period of a cyclic fluctuation would be virtually undetectable, while covariability involving only a few cycles would be detected only if it were very strong.

In spite of the fact that they are not likely to contribute a significant result, low frequency signals such as linear trends can dominate computed correlations between two time series. This can obscure relationships on shorter time scales. For example, if there were a true inverse or negative relationship between two populations on rapidly varying time scales, yet both were declining slowly for other reasons, the computed correlation between them could be positive. The problem with this result is that the computed correlation is due to variability on an essentially undetectable time scale (i.e., a linear trend), while the true correlation remains undetected. To prevent variability on undetectable time scales from occluding variability on detectable time scales, linear trends are frequently removed from time series before correlations are computed (e.g., Botsford and Kope 1992).

Another means of removing low-frequency variability from time series is first differencing. This removes low frequencies by effectively high-pass filtering the series (e.g., Thompson and Page 1989). The problem with this approach is that the consequent computation of correlations between series then is focused on covariability of the first differences of the two variables, when first differences may not be biologically meaningful. For example, the first difference of an abundance time series is a confounded combination of mortality and recruitment. For this reason, we do not employ first differencing here.

Columbia River Data

We analyze data for spring/summer, yearling migrant (stream-type) chinook salmon from the following areas in this analysis (Fig. 1):

- John Day basin (Middle Fork, North Fork/Granite, and Upper Mainstem).
- Bear Valley/Elk Creek, Marsh Creek and Sulphur Creek in the Middle Fork of the Salmon.
- Wind and Klickitat, above Bonneville Dam but below other impoundments.

- Deschutes/Warm Springs, above two impoundments.
- Entiat, Wenatchee, and Methow subbasin in the mid-Columbia; and
- Poverty Flat and Johnson Creek summer chinook in the South Fork of the Salmon River.
- Minam River spring chinook.
- Imnaha River spring/summer chinook.

We group the seven Snake stocks (Middle Fork Salmon, Poverty Flat/Johnson, Minam, and Imnaha) into a single sub-basin. We treat the mid-Columbia and the John Day streams each as separate sub-basins, and group Deschutes/Warm Springs, Wind and Klickitat together as "lower-river stocks".

The data result from spawner at age estimates based on redd counts and subsampling for age determination from scales (Petrosky, et al. 1997). These data are based on a number of assumptions and poorly known parameters such as peak index area redd count, total spawning area/index area, spawners/redd, and pre-spawning mortality, none of which are known with complete certainty. From these data we computed several different variables with different statistical and biological characteristics. The natural logarithm of the number of recruits resulting from each spawner is a commonly used population characteristic in salmon. To remove potential effects of density-dependence from this relationship, we also computed the residuals about a regression based on the Ricker stock-recruitment relationship (i.e., ln(R/S) regressed on S). Lastly, to determine whether the analyses done here could be done using spawning abundance alone (i.e., without age structure from each year), we used the ratio of spawners in one year to spawners 4 years earlier. This would be a rough approximation of recruits per spawner since ages at spawning vary from approximately one-third at age 4 and two-thirds at age 5 on Salmon River steams to 75% age 4 and 25% age 5 in the John Day streams.

The different types of population data have different statistical characteristics (Fig. 3). The most direct data are annual estimates of spawner abundance, which are based on redd counts in index streams. The spawner data is the most direct reflection of the population status because it does not depend on the assumptions involved in ageing and run reconstruction. However, it is inherently limited. Because it is the sum of several random variables, i.e., of several year classes, it contains a high degree of intra-series correlation. The residuals from detrending spawning abundance will have less intra-series correlation, hence will be a useful indicator of covariability on more detectable time scales. The natural logarithm of the number of recruits generated per spawner generating them (i.e., ln(R/S)), will be even more useful for detecting covariability because it inherently involves a natural form of differencing (ln R - ln S). It will have a higher number of degrees of freedom, but will depend to some degree on the assumptions involved in ageing and run reconstruction. The residuals from a fit to a Ricker stock-recruitment model will have a similar, high number of degrees of freedom. The ratio of spawners at one time to those A years earlier, where A is the age at which most of the spawning occurs, may approximate this variable, and would be available for populations for which run reconstructions were not available. Here we approximate that variable with the ratio of spawners at one time to spawners four years earlier.

The slow variability in some of these series limit the results that can be expected from analyses of covariability between the population time series. Analyses of abundance time series may show common declines in abundance, but it could be difficult to demonstrate that these are statistically significant declines because they are on slowly varying, marginally detectable time scales. Residuals about linear trends in abundance will be good indicators of covariability on more rapidly varying and hence more detectable time scales. Analyses of the time series reflecting survival, ln (R/S) and S_{t+4}/S_t, will vary more rapidly than abundance, hence will be useful indicators of covariability on detectable time scales. However, since a slow decline in survival is still unlikely to be detectable, a significant result will require changes on rapidly varying time scales.

Estimating Effective Degrees of Freedom

To test for significant correlations between the various series, we would ideally adjust the level of correlation required for a significant result, based on computations from equation (1) or equation (2). The problem with this approach is that the true number of degrees of freedom is difficult to estimate. While we can readily discuss the effects of the number of degrees of freedom in idealized time series (e.g., straight line trends, cycles, moving averages over various periods) empirical determination of the number of degrees of freedom in an actual time series is difficult. The various methods for estimating the standard error of computed correlations (i.e., Bayley and Hammersley 1946, Chelton 1983, 1984, Drinkwater and Myers 1987, Kope and Botsford 1990.) have the fundamental limitation that they depend on the data in each series (i.e., a single realization of each process). We do not know the values of the true correlations ρ_{xx}, ρ_{yy}, and ρ_{xy} at various lags in equations (1) and (2), hence must estimate them from these data. This inevitably leads to error in the estimate of the standard error of the estimate of the correlation coefficient or equivalently, the effective number of degrees of freedom.

For the chinook salmon on the Columbia River, we take a different approach. We have 16 series of each different type of data (e. g. 1/stock), hence we can estimate the effective number of degrees of freedom using several series of each data type, then averaging them to obtain a generic number of degrees of freedom for that type. To the degree that the different series within each data type are similar in terms of intra-series correlation, this gives us a better estimate of effective degrees of freedom than that based on a single series.

We estimated the effective number of degrees of freedom using a modified version of equation (2), which is less susceptible to variability in the estimates of the values of ρ_{xx} and ρ_{yy} from the limited data. Because higher values of the index of summation involve values of ρ_{xx} and ρ_{yy} at higher lags, for which estimates will be less precise because they are based on fewer data points, we limited the summation to a number less than N. We chose the limit of the summation based on performance of each using simulated data. We chose the value that resulted in a probability of Type I error closest to the specified error (we used 0.05).

In addition to various limits on the summation, we evaluated two different approaches to determining effective degrees of freedom for each pair of series. In one method we computed the effective degrees of freedom of each single series, then conservatively chose the smaller number of degrees of freedom (i.e., using ρ_{xx}^2 in equation (2)). In the second method we computed the effective degrees of freedom from pairs of series (i.e., using $\rho_{xx}\rho_{yy}$ in equation (2)).

The results of these trials indicate summing up to j=10 in equation (2) using pairs of series, rather than single series give reasonable probabilities of type I error (Table 1). Single series estimates over-estimate the variance leading to an under-estimate of the effective degrees of freedom, a high significance threshold and a low probability of Type I error. Of the paired series estimates, the variance is underestimated at j=5 and overestimated at j=30, with the consequent expected changes in probability of Type I error. Whether the series are Gaussian or Uniformly distributed appears to make little difference for these results.

Based on these results, we computed the effective degrees of freedom for each data type (spawners, ln (R/S), etc.) by calculating the variance from equation (2) for each of 63 pairwise combinations (i.e., there are 63 possible pairwise correlations for the indicators for 16 stocks). We did not allow the effective degrees of freedom to exceed the series length N. We then use the 63 N*'s to calculate an average effective degrees of freedom (EDF) for that data type. We then compare this average N* to the average, uncorrected degrees of freedom, and calculate a ratio of average N* to average uncorrected N. To estimate the EDF for a given pairwise correlation, we next calculate EDF(x,y) as (average N*/average N) * (N(x,y)). This EDF(x,y) is then used to calculate the significance of each correlation.

Results

Although the EDFs calculated from single series were biased low, they are our only source of information for comparison of the relative EDF for the different streams (Table 2). Some of the Snake River streams had low EDFs as might be expected from the dominance of a linear trend. Note the low EDF for Bear Valley/Elk due to the dominant declining trend in Fig. 3. In all but two cases, detrending increased the EDFs, as would be expected. In more than half the cases, EDFs for logarithm of recruits per spawner was greater than that for spawners. A striking exception is the Wenatchee.

The EDF's computed from pairs of series range from 18.5, for Ricker regression residuals, to 25.5, for detrended spawners (Table 3). In all cases, de-trending increased the EDF, as expected. It is somewhat surprising that the EDF of spawning escapement is roughly the same as (actually slightly larger than) EDFs of the other types of series. Since we had expected that spawning escapement would have a lower EDF, we recomputed the mean group EDF for the Snake River subbasin only, to see whether the other streams had inflated that EDF. This resulted in an EDF of 13.3, which is much less than the value of 20.4 in Table 3, but slightly greater than the average value of the single series estimates in Table 2, 11.1.

The spatial pattern of covariability indicated by the spawning escapement data is moderately consistent covariability between stocks within the three sub-basins, with some covariability between the Snake River sub-basin and the Mid Columbia (Tables 4, 8). Within sub-basins there are fewer significant correlations in the detrended data, due to a decline in correlation. A decline in correlation between a raw series and a detrended series indicates the relationship in the trends is the same as the relationship at higher frequencies, in this case a positive correlation. The correlation between the Klickitat river and some of the Snake River stocks, which is stronger in the detrended data, is notable because of the large distance between the stocks (Fig. 1).

The spatial pattern in the logarithm of recruits per spawner is stronger than that in spawning escapement (Tables 5, 8). Correlations are uniformly higher, and more are significant. The strongest relationships are within sub-basins with all streams significantly correlated with all other index streams in their sub-basin for all three sub-basins. Within sub-basins correlations in the detrended data are again less in all cases except Poverty Flat. The significant relationships between Klickitat River and Snake River sub-basin stocks is again stronger in the detrended data.

The spatial pattern in the Ricker residuals is similar to that in the logarithm of recruits per spawner except that there are fewer significant correlations between sub-basins (Tables 6, 8). In the correlations within sub-basins, the values of correlation for the detrended data are almost the same as the corresponding values for data without the trends removed. This is probably due to the fact that the trends in these data are less. Covariability between the Klickitat and the Snake River sub-basins is present in this data type also.

The spatial pattern in the ratio of spawning escapement in one year to spawning escapement four years earlier differs from that of the raw spawning escapement data in that the correlations within each sub-basin are stronger (Tables 7, 8). It reflects the pattern in the logarithm of recruits per spawner and the Ricker residuals better than the spawning escapement data. A possible drawback to its use as a replacement for recruitment data is the fact that it indicates more between sub-basin relationships than the types of data involving recruitment.

In all of these types of data, the strongest inter-sub-basin relationships seem to be between the Snake River sub-basin, and the Mid Columbia, with a lesser relationship with the John Day. The Mid Columbia and the John Day are related to each other but not as strongly. These indications of inter-sub-basin relationships are weaker in the higher frequency series (i.e., the series with trends removed, and the Ricker residuals).

Discussion

The dominant characteristic of the spatial pattern among these Columbia River spring/summer chinook salmon stocks is covariability between survival indicators within the three sub-basins. This pattern is strongest in the three types of data directly reflecting survival, but is clearest in the Ricker residuals. The fact that this appears in data reflecting survival rather than abundance is not surprising. Survival varies from year to year, while abundance is the cumulation of many random past survivals. Covariability between sub-basins is only moderate, but is strongest between the Snake river sub-basin and the Mid-Columbia. The large distance between the Klickitat River and the Snake River sub-basin suggest the unusual covariability between these stocks may be spurious.

This analysis demonstrates the value of accounting for intra-series correlation. The level at which a correlation coefficient would have been termed significant for most of these series if intra-series correlation were not accounted for is roughly 0.3 (Fig. 2, N=30). With this as a criterion, one can readily see that many more correlations would have been identified as significant, thus blurring the observed spatial patterns. Also, if a more stringent criterion had been applied to spawner abundance data from the Snake River stocks, as may be appropriate to that data type with declining trends, even fewer relationships among spawner data would have been identified as significant.

The similarity between the spatial pattern of covariability indicated by the analysis of spawners in one year divided by spawners 4 years earlier is encouraging. Note the similarity in time series in Fig. 3. Since there are spawner abundance data available for many more stocks for which recruitment data (or ageing) is not available, analyses similar to this one can be performed over a wider scale. This ratio successfully detects covariability of survival reflected in logarithm of recruits per spawner and Ricker residuals, but appears to be less selective than those data types. In our case, it may have been influenced by the slight variability in dominant age of spawning among Columbia Basin stocks.

Although we have carefully accounted for effective degrees of freedom so that we could weigh the relative importance of correlations with different types of data, the overall number of degrees of freedom is much lower than the figures used due to a complete lack of accounting for multiplicity of tests. The results described here should not be considered a formal hypothesis test in the sense that we have statistically proven that any of the results did not arise from chance alone. We consciously sacrificed the ability to do formal hypothesis testing for the sake of searching for pattern among as wide a suite of geographical locations and data types as possible. The next step would be testing these relationships through other means.

References

Bayley, G.V. and J.M. Hammersley. 1946. The effective number of independent observations in an autocorrelated series. Journal of the Royal Statistical Society B8: 184-197.

Bence, J.R. 1995. Analysis of short time series: correcting for autocorrelation. Ecology 76: 628-639.

Botsford, L. W. and J.G. Brittnacher. 1992. Detection of environmental influences on wildlife: California quail as an example. Pp. 158-169 in D. R. McCulloch and R.H. Barrett (eds.). Elsevier Applied Science. 1163pp.

Botsford, L.W. and T.C. Wainwright. 1996. Detection of environmental influences on biological populations.

Campbell, A. and R.K. Mohn. 1983. Definition of American lobster stocks for the Canadian Maritimes by analysis of fishery landing trends. Trans. Amer. Fish. Soc. 112: 744-759.

Chelton, D.B. 1983. Effects of sampling errors in statistical estimation. Deep-Sea Research 30: 1083-1103.

Chelton, D.B. 1984. commentary: short-term climatic variability in the Northeast Pacific Ocean. Pp. 87- 99 in W. Pearcy, ed. Aspects of climate variability in the Pacific and the Western Americas. Oregon State University press, Corvallis Oregon, USA.

Cohen, E.B., D.G. Mountain and R. O'Boyle. 1991. Local-scale versus large scale factors affecting recruitment. Can. J. Fish. Aquat. Sci. 48: 1003-1006.

Drinkwater, K.F. and R.A. Myers. 1987. Testing predictions of marine fish and shellfish landings from environmental variables. Can. J. Fish. Aquatic Sci. 44: 1568-1573.

Elner, R. W. and A. Campbell. 1991. Spatial and temporal patterns in recruitment for American lobster, *Homarus americanus*, in the northeastern Atlantic. Mem Queensl. Mus. 31: 349-363.

Kope, R.G. and L.W. Botsford. 1990. Determination of factors affecting recruitment of chinook salmon *Oncorhynchus tshawytscha* in Central California. U.S. Fish. Bull. 88: 257-269.

Koslow, A.J., Thompson, K.R. and Silvert, W. 1987. Recruitment to northwest Atlantic cod (*Gadus morhua*) and haddock (*Melanogrammus aeglefinus*) stocks: influence of stock size and climate. Can. J. Fish. Aquat. Sci., 44: 26-39.

Petrosky, C E, H A Schaller and R.C. P. Beamesderfer. 1995. Spawner-recruit relationships for spring and summer chinook salmon populations in several Columbia and Snake River subbasins. ms.

Thompson, K R and F.H. Page. 1989. Detecting synchrony of recruitment using short, autocorrelated time series. Can. J. Fish. Aquat. Sci. 46: 1831-1838.

Table 1. Performance of various methods of correcting for intraseries correlation. Mean Effective Degrees of Freedom (EDF) and actual probability of Type I error for a series with no intraseries correlation (i.e. white) and a four-point moving average generated from a Gaussian and a Uniform Distribution

			Gaussian		Uniform	
			White Series	MA4 Series	White Series	MA4 Series
	Uncorrected	**Mean EDF**	36	36	36	36
		Probability	0.050	0.295	0.045	0.258
j=5	*Pairwise Correction*	**Mean EDF**	36.03	18.83	35.99	18.79
		Probability	0.053	0.134	0.045	0.121
	Single Series Correction	**Mean EDF**	34.41	12.96	34.38	12.97
		Probability	0.057	0.131	0.049	0.166
j=10	*Pairwise Correction*	**Mean EDF**	36.58	12.16	36.38	12.05
		Probability	0.057	0.039	0.047	0.047
	Single Series Correction	**Mean EDF**	28.95	9.34	28.84	9.31
		Probability	0.023	0.058	0.025	0.071
j=30	*Pairwise Correction*	**Mean EDF**	41.13	9.32	39.79	9.22
		Probability	0.068	0.015	0.058	0.016
	Single Series Correction	**Mean EDF**	15.04	3.62	14.88	3.63
		Probability	<0.001	<0.001	<0.001	<0.001

Table 2. Actual number of data points in spawning escapement and effective degrees of freedom (EDF) of various data types estimated from the single series version of Eq (2).

Location	Subbasin	Years of Data	Spawners	Spawners Detrended	Ln (R/S)	ln(R/S) Detrended
Snake						
	Bear Valley/Elk	32	6.09	22.38	15.75	16.83
	Sulphur Creek	32	14.38	20.92	18.36	19.95
	Marsh Creek	32	8.77	20.01	15.25	16.97
	Minam	32	16.84	25.73	13.66	13.92
	Imnaha	32	17.73	24.73	16.71	17.88
	Johnson	32	8.41	25.26	22.04	21.53
	Poverty Flat	32	4.97	12.32	20.2	20.22
Mid-Col.						
	Entiat	32	18.81	21.85	17.09	17.88
	Methow	31	15.35	22.15	21.22	21.13
	Wenatchee	32	24.12	23.85	7.68	23.06
BONN-MCN						
	Upper John Day	32	19.77	22.86	23.95	29.88
	Middle Fork John Day	32	16.99	18.18	22.43	21.99
	North Fork John Day	32	17.52	18.62	16.29	20.03
BONN-MCN						
	Warm Springs	22	9.4	10.09	9.18	10.11
	Klickitat	25	19.25	19.23	12.01	12.02
	Wind	21	15.35	16.31	12.95	14.3

Figure 1. A map of streams.

Figure 2

. The threshold value a correlation coefficient must have to be significant at each of three levels for various effective degrees of freedom in the time series.

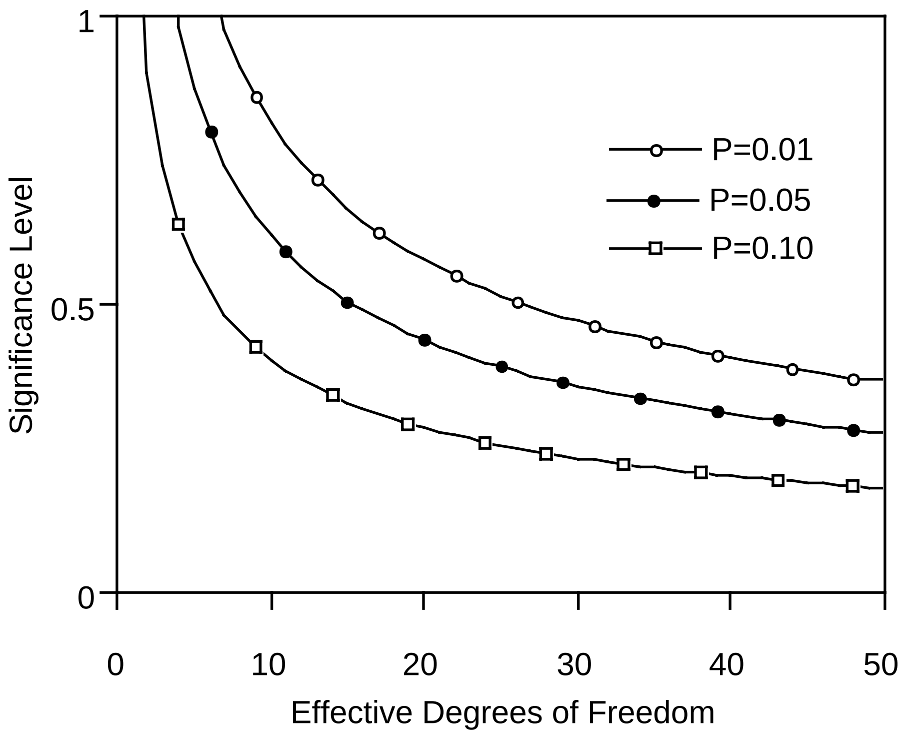

Figure 3. An example of different data types from an upriver index stream and a lower river index stream.